Research Note

UNITED STATES DEPARTMENT OF AGRICULTURE
FOREST SERVICE
INTERMOUNTAIN FOREST & RANGE EXPERIMENT STATION
OGDEN UTAH

No. 45 May 1957

HOW MANY TREES SHOULD BE LOG GRADED TO DETERMINE SALES REALIZATION VALUE[1]

Donald W. Lynch
Division of Forest Management Research

With the increased use of log grading in timber sale appraisals, foresters are concerned about the accuracy of the sales realization value that they determine from grading a portion of the trees in a timber sale. Grading obviously must be accomplished by sampling, since the job would be too large on most sales to grade every tree. But how large must the sample be? As in any sampling problem, the variation of the units being measured, the desired accuracy of the results, and the size of the total population of units (in this case, the number of trees in the sale) are the factors governing sample size.

The sample accuracy (or allowable sampling error) is set by policy, and total number of trees in the sale is estimated from the cruise. There remains, then, the variation of the sampling units to determine. The sampling unit in this case is the realization value, per thousand board feet, of a single tree as determined by the value of each log in the tree.

To obtain an estimate of the variation of tree values for a given sale, a preliminary sample must be obtained. Twenty to thirty trees, representative of the population, should be graded, measured, and their value per thousand board feet, "y," determined as follows:

1/ Sales realization is defined as the gross sales price of manufactured lumber. It is currently adjusted and used as a basis for stumpage appraisals.

$$y = \frac{V_1P_1 + V_2P_2 + \ldots}{V_1 + V_2 + \ldots}$$

where V_1, V_2, etc. = Volume in board feet <u>lumber scale</u> of logs in the tree

P_1, P_2, etc. = prices per M of lumber in the particular logs of the tree

From the twenty to thirty observations of "y" thus obtained, the standard deviation and coefficient of variation are computed. Then for a given accuracy and a given total number of trees, the sample size needed can be calculated readily. (See appendix for formulae and numerical example.)

Knowing the required sample size, the additional trees should be graded and their values computed and combined with the preliminary sample to obtain the <u>average realization value</u> for a particular timber sale. This average will be within the desired accuracy at the established level of probability if the sample trees have been chosen in an unbiased and random fashion.

Table 1 shows the sample size needed for accuracies of 0.05, 1, 2, and 3 percent, for coefficients of variation of 0.05, 0.10, 0.15, and 0.20, and for total sale sizes of 500, 1,000, 2,000, 5,000, and 10,000 trees, all at the 95 percent level of probability.

Other values can be determined from the table by plotting on cross-section paper the tabular values of sample size over one of the variables (coefficient of variation, accuracy, or total number of trees) while holding constant the other two variables. Fit a freehand curve through the plotted points and read interpolated values from the curve.

Table 1.--Number of trees, "n," needed in a sample for various accuracies, coefficients of variation, and total sale sizes, computed at the 95 percent probability level

Total sale size	CV 0.05				CV 0.10				CV 0.15				CV 0.20			
	Accuracy percent				Accuracy percent				Accuracy percent				Accuracy percent			
	1/2	1	2	3	1/2	1	2	3	1/2	1	2	3	1/2	1	2	3
(trees)	(sample size, "n")															
500	222	83	24	11	381	222	83	41	439	321	155	83	464	381	222	131
1,000	286	91	24	11	615	286	91	43	783	474	184	91	865	615	286	151
2,000	333	95	25	11	889	333	95	43	1,285	621	202	95	1,524	889	333	163
5,000	370	98	25	11	1,212	370	98	44	2,093	763	215	98	2,807	1,212	370	172
10,000	385	99	25	11	1,379	385	99	44	2,647	826	220	99	3,902	1,379	385	175

APPENDIX

Formula for the standard deviation (in a form convenient for machine calculation):

$$s_y = \sqrt{\frac{\overset{n}{S}(y^2) - \frac{\left[\overset{n}{S}(y)\right]^2}{n}}{n - 1}}$$

where s_y = standard deviation of y

y = an individual sampling unit

n = the number of sampling units in the sample

$\overset{n}{S}$ = "the sum of n units"

Formula for the coefficient of variation:

$$CV = \frac{s_y}{\bar{y}}$$

where CV = coefficient of variation

$\bar{y}$ = mean of the sample of y's

Formula for the required number of sampling units:

$$n = \frac{t^2(CV)^2}{E^2 + \frac{t^2(CV)^2}{N}}$$

where n = required number of sampling units

t = standard units. (a t of 1 for approximately 68 percent probability and a t of 2 for approximately 95 percent)

E = desired accuracy in decimal form

N = total number of sampling units in the population (in this case, total number of trees in the sale)

Numerical example:[2]

Logs in tree No. 1: 21" grade 5, 19" grade 5, 17" grade 5, 14" grade 6.

$$y_1 = \frac{(31)(89.52)+(26)(90.36) + (21)(90.01)+(14)(78.61)}{31 + 26 + 21 + 14}$$

$$= 88.21$$

Logs in tree No. 2: 30" grade 1, 28" grade 2, 26" grade 5, 23" grade 5, 19" grade 5, 15" grade 5.

$$y_2 = \frac{(64)(155.76)+(56)(129.72) + (48)(91.37)+(38)(90.36)+(26)(90.01) + (16)(92.22)}{64 + 56 + 48 + 38 + 26 + 16}$$

$$= 116.40$$

Continue in a similar manner until 20 to 30 values of y are computed.

2/ Grateful acknowledgment is given to the Boise National Forest for the use of these data which were taken from an actual timber sale of ponderosa pine.

y[3]
88.2
116.4
91.7
114.0
89.1
110.8
94.5
103.0
117.4
109.5

$\overset{n}{S}(y) = 1034.6$

$\bar{y} = 103.46$

Computing the standard deviation:

$\overset{n}{S}(y^2) = 108{,}260.80$

$\dfrac{\left[\overset{n}{S}(y)\right]^2}{n} = 107{,}039.72$

Difference $= 1{,}221.08$

Dividing by $(n-1)$ $= 135.68$

Extracting square root $= 11.65 = s_y$

Computing the coefficient of variation:

$$CV = \frac{11.65}{103.46} = 0.11$$

Computing required sample size:

(given: desired accuracy of 2 percent, total number of trees 1200, and $t = 2$ for a 95 percent probability)

$$n = \frac{(2)^2(.11)^2}{(.02)^2 + \dfrac{(2)^2(.11)^2}{1200}} = \frac{.0484}{.00044} = 110$$

3/ For purposes of this example, only ten values of y are included. Normally, 20 to 30 such values should be used to get a reliable estimate of the coefficient of variation. Values have been rounded to the nearest tenth of a dollar.